

**UNIVERSIDAD AUTONOMA DE MADRID**

**ESCUELA POLITECNICA SUPERIOR**



**TRABAJO FIN DE MÁSTER**

**Análisis de interacciones de letalidad  
sintética (SL) en cáncer y predicción de  
tratamientos**

Máster Universitario en Bioinformática y Biología  
Computacional

**Autor:** FUSTERO TORRE, Coral

**Tutor:** Al-Shahrour Núñez, Fátima  
Unidad de Bioinformática,  
Centro Nacional de Investigaciones Oncológicas

**FECHA:** Febrero, 2018



# **IDENTIFICACIÓN *IN SILICO* DE NUEVAS DIANAS TERAPÉUTICAS EN CÁNCER BASADAS EN *LETALIDAD SINTÉTICA*.**

Autora: Coral Fustero Torre

Tutora: Fátima Al-Shahrour Núñez

Ponente: Ana González Marcos

Unidad de Bioinformática

Centro Nacional de Investigaciones Oncológicas

**Escuela Politécnica Superior**

**Universidad Autónoma de Madrid**

**Febrero, 2018**



# Índice

## 1. Resumen

## 2. Objetivos del proyecto

## 3. Introducción

### 3.1. Cáncer y medicina personalizada

#### 3.1.1. Identificación de dianas terapéuticas

### 3.2. Retos de las terapias dirigidas

### 3.3. Letalidad sintética en cáncer

#### 3.3.1. Clases de letalidad sintética

#### 3.3.2. Retos de la técnica

#### 3.3.3. Detecciones de letalidad sintética en células humanas

##### 3.3.3.1. Detecciones mediante ARN de interferencia

##### 3.3.3.2. Detecciones mediante CRISPR-Cas9

##### 3.3.3.3. ARNi vs. CRISPR-Cas9

#### 3.3.4. De la detección de interacciones de letalidad sintética a la práctica clínica

### 3.4. Variantes genómicas

## 4. Materiales y métodos

### 4.1. Conjuntos de datos

#### 4.1.1. SynLethDB

#### 4.1.2. PanDrugs

#### 4.1.3. Project Achilles

#### 4.1.4. Cancer Cell Line Encyclopedia

### 4.2. Integración de datos

#### 4.2.1. Tratamiento de los datos

#### 4.2.2. Paso 1: Cruce entre Pandrugs y SynLethDB

#### 4.2.3. Paso 2: Filtrado de los pares de SL de interés mediante el uso de información genómica

##### 4.2.3.1. Para la información de mutaciones puntuales

##### 4.2.3.2. Para el análisis de alteraciones en número de copia

##### 4.2.3.3. Consenso

#### 4.2.4. Paso 3: Análisis de viabilidad celular

#### 4.2.5. Paso 4: Análisis estadístico

#### 4.2.6. Paso 5: Caracterización y anotación

### **5. Resultados**

#### 5.1. Resultados derivados del filtrado

##### 5.1.1. Genes con terapias directas seleccionadas

##### 5.1.2. Fuentes de origen de los datos

#### 5.2. Anotación

#### 5.3. Tabla de resultados

#### 5.4. Casos de estudio

##### 5.4.1. Caso 1: BRCA1 – PARP1

##### 5.4.2. Caso 2: ERBB2 – PARP1

##### 5.4.3. Caso 3: KRAS – UGCG

### **6. Discusión**

### **7. Conclusiones**

### **8. Bibliografía**

### **9. Anexos**

## 1. Abstract

The use of targeted therapies still represents challenge in Cancer research. With the discovery of the phenomenon known as oncogenic addiction, based on the concept that there are tumours which survival depends on certain genes or biological pathways, a paradigm shift occurred on the field of drug therapies research and development, now focused on *personalized medicine*.

The idea behind personalized medicine results from the dependency between the therapy efficacy and the presence of specific genetic alterations. And its main goal is the development of drugs targeting the tumour's specific mutations. But even though personalized medicine is one of the most important lines of research in cancer treatment, so far this approach has only been valid in very specific tumour types and most of the patients end up resorting to traditional therapies.

One of the most promising strategies for the improvement and development of anticancer therapies is based on the use of secondary dependencies, not necessarily oncogenic, known as synthetic lethality interactions (SL). It is known that cancer cells have compensatory mechanisms that help them survive when they accumulate mutations in critical genes, developing dependencies on these genes in most of the cases. The concept of synthetic lethality uses these dependencies in order to affect the tumour cell viability.

Using large cancer cell lines screenings, some research groups have explored the effects of gene silencing and its translation into phenotype, detecting cancer cell dependencies.

In this project, we have integrated information coming from different databases focused on the recollection of synthetic lethal relationships, together with point mutations information, copy number alterations as well as gene silencing data. As a result, we have been able to identify new targets and biomarkers that predict cancer response. Specifically, we have started with more than 16 thousand synthetic lethality interactions and analysed those with the therapeutic options available (34% of the total). Exploring these interactions, we have obtained new strategies concerning undruggable oncogenes, such as KRAS, or tumour suppressor genes, such as BRCA. Also, the analysis has helped us delve into the biological mechanisms behind these interactions.

In order to properly characterize the identified interactions, it will be necessary to add more layers of information to our approach, such as expression data and methylation, as well as exploring the pharmaceutical inhibition of these targets.

Adding synthetic lethality to the already existing strategies, we will be a step closer to the classification of cancer subtypes and its subsequent treatment using targeted therapies.

**Key words**

Synthetic lethality, personalized medicine, targeted therapies, undruggable genes



## 1. Resumen

El tratamiento de tumores de forma dirigida todavía supone un desafío en la investigación contra el cáncer. Con el descubrimiento del fenómeno conocido como adicción oncogénica, basado en el concepto de que existen tumores cuya supervivencia depende de (o son adictos a) ciertos genes o rutas biológicas, se produjo un cambio de paradigma en el campo del desarrollo de fármacos antineoplásicos, ahora más dirigido hacia lo que conocemos como *medicina personalizada*.

La idea de *medicina personalizada* surge de la dependencia que existe entre la eficacia de una terapia y la presencia de determinadas alteraciones genéticas, y busca desarrollar fármacos que actúen sobre dianas específicas del tumor. Pero pese a ser una de las líneas de investigación más importantes para el tratamiento de la enfermedad, de momento sólo ha demostrado ser efectiva en tumores muy determinados y la mayoría de pacientes tienen que recurrir a terapias tradicionales.

Una de las estrategias más prometedoras para la mejora y el desarrollo de terapias anti-cáncer está basada en la explotación de dependencias secundarias, no necesariamente oncogénicas, conocidas como interacciones de letalidad sintética (SL).

Se sabe que las células cancerígenas tienen mecanismos de compensación que les ayudan a sobrevivir en el caso de acumular mutaciones en genes críticos, desarrollando dependencias para con dichos genes en muchos de estos casos. El concepto de letalidad sintética surge con la finalidad de aprovechar estas compensaciones y así afectar la viabilidad de la célula tumoral.

Utilizando pruebas de detección masivas a lo largo de paneles de líneas celulares tumorales, algunos grupos de investigación han explorado los efectos del silenciamiento de genes y su relación con el comportamiento fenotípico de la célula, detectando dependencias en células cancerígenas.

En este proyecto, hemos integrado la información recogida en diversas bases de datos dedicadas a la recolección de interacciones de letalidad sintética, con información de mutaciones puntuales, alteraciones en número de copia y silenciamiento génico.

A partir de este análisis se han podido identificar nuevas dianas terapéuticas y biomarcadores predictivos de respuesta en cáncer. En concreto, hemos partido de más de 16 mil interacciones de letalidad sintética y analizado aquellas con opciones terapéuticas disponibles (un 34% del total).

La exploración de estas interacciones nos ha llevado a la obtención de nuevas estrategias terapéuticas que involucran a oncogenes sin disponibilidad de tratamientos dirigidos como KRAS o genes supresores de tumores como BRCA. Además, nos ha ayudado a profundizar en los mecanismos biológicos que se encuentran detrás de estos eventos.

Para completar ese análisis será necesario integrar más fuentes de información como datos de expresión, metilación, así como explorar los efectos de la inhibición farmacológica de estas dianas. Añadiendo la estrategia de letalidad sintética a las ya existentes, podemos avanzar en la definición de los subtipos de cáncer y su tratamiento mediante terapias dirigidas.

## **Palabras clave**

*Letalidad sintética, medicina personalizada, terapias dirigidas, genes undruggable*

## 2. Objetivos del proyecto

El objetivo principal del proyecto consiste en la identificación de nuevas dianas terapéuticas y biomarcadores predictivos de respuesta al tratamiento en cáncer empleando para ello el concepto de Letalidad Sintética (SL).

Los objetivos específicos son:

- Procesamiento de los datos provenientes de SynLethDB (Jing Guo *et al.*, 2016), un recurso dedicado a la recolección de pares de SL provenientes de otras bases de datos, ensayos bioquímicos, predicciones computacionales y minería de datos.
- Filtrado de estos datos según las terapias disponibles en la herramienta PanDrugs (<http://www.pandrugs.org>), un método para la selección de terapias a partir del análisis de alteraciones moleculares presentes en cáncer.
- Análisis de la información genómica (en nuestro caso, número de copia y mutaciones puntuales) disponible para los pares de interacción seleccionados en el catálogo de 501 líneas celulares pertenecientes al proyecto Achilles (*Project Achilles*), destinado a generar un catálogo de esencialidad génica para cientos de líneas celulares relacionadas con cáncer.
- Integración con información de viabilidad celular derivada del proyecto Achilles.
- Caracterización de los posibles candidatos terapéuticos detectados.

Como objetivo último, se buscará integrar la información ya curada a la plataforma PanDrugs, una guía para la selección de terapias a partir de resultados de genómicos en cáncer, desarrollada por el grupo de investigación de Fátima Al-Shahrour, añadiendo posibles nuevas dianas efectivas y ampliando su espectro de acción.



### 3. Introducción

El cáncer se trata de un conjunto de enfermedades caracterizadas por el desarrollo de células anormales, que se dividen de forma incontrolable y tienen la capacidad de infiltrarse y destruir el tejido normal. Es una de las principales causas de mortandad a nivel mundial, y se estima que mata a más de 8 millones de personas cada año (Ferlay J *et al.*, 2013) .

El cáncer es causado por cambios en el genoma de las células. Estos cambios pueden ser provocados tanto por **agentes externos**: como radiación, agentes químicos, contaminación o agentes infecciosos (como el virus del papiloma humano o el virus de la hepatitis B); por **alteraciones genéticas**: adquiridas durante la replicación del ADN o heredadas; o bien, producto de **mecanismos de regulación epigenéticos**: como metilación o microARNs.

#### 3.1. Cáncer y medicina personalizada.

El desarrollo de terapias contra el cáncer depende de nuestro progreso en el entendimiento de los procesos biológicos que se encuentran detrás de la enfermedad. Actualmente, el tratamiento del cáncer está basado principalmente en tres estrategias: cirugía, quimioterapia y radioterapia; pero la complejidad de esta enfermedad hace necesaria la mejora de los tratamientos disponibles.

Con el descubrimiento del fenómeno conocido como **adicción oncogénica** (en inglés, *oncogene addiction*), basado en el concepto de que existen tumores cuya supervivencia depende de (o son adictos a) ciertos genes o rutas biológicas, se produjo un cambio de paradigma en el campo del desarrollo de fármacos antineoplásicos, ahora más dirigido hacia lo que conocemos como **medicina personalizada** (Brunen D, Bernards R, 2017).

La idea de *medicina personalizada* surge de la **dependencia** que existe **entre la eficacia de una terapia y la presencia de determinadas alteraciones genéticas**. Se busca desarrollar fármacos que actúen sobre dianas específicas del tumor. De esta forma, no sólo estaríamos personalizando el tratamiento según las características de cada paciente, sino que se conseguiría un aumento de la ventana terapéutica disponible.

### 3.1.1. *Identificación de dianas terapéuticas.*

El desarrollo de terapias dirigidas requiere la identificación de dianas terapéuticas, es decir, proteínas nativas cuya actividad se ve modificada mediante la acción de una droga, resultando en un cambio en el funcionamiento y comportamiento de la célula.

Generalmente, estas dianas estarán relacionadas con vías que controlan el crecimiento y supervivencia de la célula.

Básicamente, la terapia dirigida buscará cambios en las células cancerígenas con respecto a las células normales e intentará contrarrestar sus efectos.

No todos los tumores tendrán las mismas dianas, por lo que la misma terapia dirigida no tiene por qué funcionar para todos los pacientes.

Un ejemplo de a dónde nos puede llevar esta estrategia, es la utilización de *imatinib* en pacientes de leucemia crónica mieloide. Imatinib fue la primera terapia dirigida aprobada por la Food and Drug Administration (FDA) de E.E.U.U. para el tratamiento de la enfermedad. La mayoría de los casos de leucemia crónica mieloide (más del 90%), están causados por una fusión entre el gen ABL (*Abelson tyrosine kinase*), que se encuentra en el cromosoma 9, y el gen BCR (*Breakpoint cluster*), situado en el cromosoma 22. *Imatinib* se encarga de inhibir esta tirosina quinasa fusionada (BRC/ABL) de forma específica y actualmente se utiliza como primera línea de tratamiento para enfermos de leucemia crónica mieloide. Gracias a esta estrategia, más del 83% de los pacientes de leucemia crónica mieloide tienen una supervivencia superior a 10 años, comparativamente con el 20% antes de su implantación (Brunen D, Bernards R, 2017). Este mismo concepto también ha sido trasladado a el tratamiento de otros tipos de cáncer, como en el caso del uso de inhibidores de EGFR en pacientes con carcinoma pulmonar no microcítico.

Cabe señalar que las mutaciones identificadas no siempre tienen por qué servir de dianas terapéuticas directas, en ocasiones se utilizan como biomarcadores. Es decir, mutaciones diferenciales entre tejido tumoral y normal, capaces de predecir o asociarse a la respuesta a determinadas drogas. Por ejemplo, en cánceres de tipo colorrectal se utilizan las mutaciones de KRAS (un oncogen que influye en el crecimiento tumoral y dispersión), presentes en el 40% de los pacientes (Dinu D *et al*, 2014) como marcador de respuesta a fármacos. Se sabe que cuando se tiene esta mutación, *cetuximab* (erbitux) y *panitumumab* (vectibix) no son efectivos, de forma que se podrá proponer un tratamiento más efectivo como primera línea de acción.

En resumen, la medicina personalizada tiene el potencial de ofrecer terapias individuales y altamente específicas con muchos menos efectos secundarios. Además, evita la aplicación

de fármacos que no tienen ningún beneficio para el paciente, así como la administración de drogas más generales y por tanto más citotóxicas, agregando un mayor nivel de protección para el paciente (Shah NP *et al.*, 2004).

### **3.2. Retos de las terapias dirigidas.**

Gracias a los avances de los últimos años en materia de secuenciación, ahora es posible detectar cambios epigenéticos y genéticos que diferencian las células tumorales de un paciente, de sus células normales. Este tipo de alteraciones genéticas específicas de tumor, revelan no sólo cambios biológicos que dirigen la progresión del tumor, sino también las vulnerabilidades que pueden ser explotadas a la hora de utilizar drogas dirigidas.

Pero pese a que las terapias dirigidas son una línea de investigación muy importante para el tratamiento del cáncer, de momento sólo han demostrado ser efectivas para tumores muy determinados, por lo que la mayoría de los pacientes tienen que recurrir a terapias tradicionales.

Existen varios factores que influyen en el progreso de esta estrategia: por un lado, la mayoría de estas terapias explotan el fenómeno de la adicción oncogénica, lo cual limita mucho el repertorio de dianas estudiadas para el desarrollo de terapias. La industria farmacéutica dedica la mayoría de sus recursos al estudio de fármacos para atacar mutaciones de ganancia de función en drivers oncogénicos, mientras que muchos cánceres están provocados por mutaciones de pérdida de función. Por otro, las terapias dirigidas están altamente limitadas por la adquisición de resistencias.

Además, se trata de una táctica que no es aplicable para todos los oncogenes y existen muchos que actualmente no tienen una terapia disponible (*undruggable*). Un ejemplo serían las proteínas oncogénicas de la familia RAS, para las que el desarrollo de fármacos que permitan su inhibición ha tenido muy poco éxito, pese a que esta familia de genes está mutada en más del 30% de los tumores y que las células tumorales dependan de su activación para sobrevivir y proliferar (Karnoub AE, Weinberg RA, 2008; Roberts PJ, Der CJ, 2007).

### 3.3. Letalidad sintética en cáncer.

Es por ello que la identificación de terapias específicas para las distintas mutaciones o patrones mutacionales sigue siendo uno de los principales retos para el desarrollo del campo de la medicina de precisión (Sinha S et al., 2017).

Una de las estrategias más prometedoras para la mejora y el desarrollo de terapias anti-cáncer está basada en la explotación de dependencias secundarias, no necesariamente oncogénicas, conocidas como interacciones de **letalidad sintética** (SL).

Definimos una interacción de letalidad sintética, como aquella que ocurre en dos genes cuando la perturbación de uno de ellos no afecta la viabilidad de la célula, mientras que la perturbación de ambos, provoca la muerte celular (O'Neil NJ et al., 2017).

La letalidad sintética se basa en la incompatibilidad entre la supervivencia celular y la pérdida de función de dos o más genes.

En el contexto de cáncer, los eventos de SL también se denominan **adicción no oncogénica** (*non-oncogene addiction*), debido a que las células mutantes requieren (o son adictas) a la actividad del par de letalidad sintética para su viabilidad.

El producto proteico de un gen que tiene una interacción de letalidad sintética con una mutación somática frecuente y específica de tumor, sería un candidato excelente para una terapia anti-cáncer. La droga propuesta explotaría una interacción que sólo afectaría a las células tumorales con dichas mutaciones y no a las células normales, reduciendo así sus posibles efectos secundarios y aumentando su sensibilidad.

Es decir, esta estrategia podría resultar en un aumento significativo de la eficacia terapéutica. Como otra de sus ventajas, las interacciones de letalidad sintética pueden expandir el repertorio de terapias anti-cáncer disponibles, gracias al reposicionamiento de fármacos (Weide UH et al., 2011).

#### 3.3.1. Clases de letalidad sintética

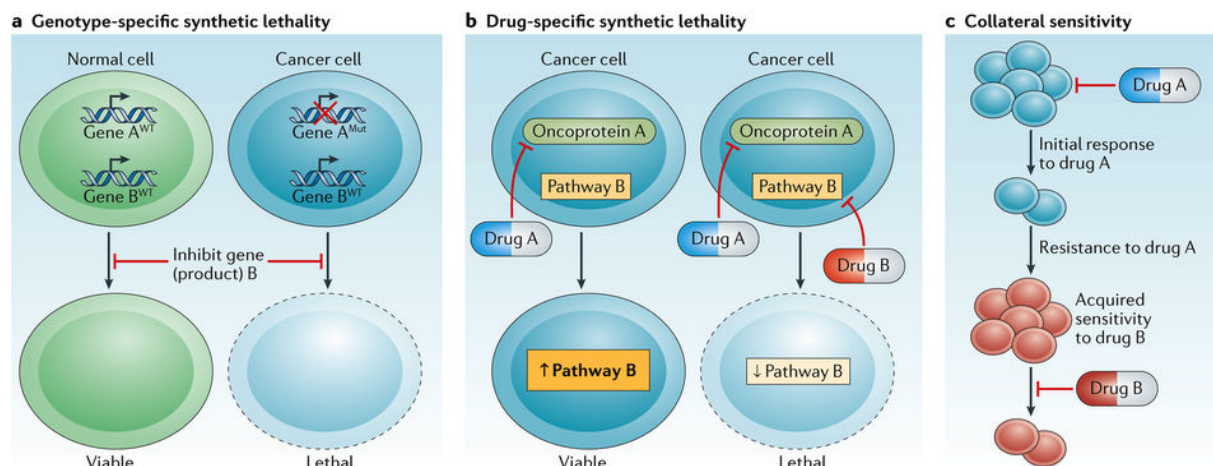
Existen diversas clasificaciones para determinar los tipos de eventos de letalidad sintética que se pueden dar en una célula. (O'Neil et al., 2017)

1. Basándonos en **eventos mutacionales**, podemos clasificar la letalidad sintética principalmente en tres subtipos.
  - a. Pérdida de función/Pérdida de función. Dado un perfil genómico o farmacológico de “inactivación”, se refiere a la inhibición de otro gen o ruta,



con el fin de inducir letalidad sintética. Un ejemplo de este tipo de perfil genómico podría referirse inactivación de p53, un gen supresor de tumores comúnmente mutado en cáncer.

- b. Ganancia de función/Pérdida de función (letalidad sintética de dosis o *synthetic dosage lethality*). Uno de los mayores contribuidores en el desarrollo de un tumor es la activación de oncogenes mediante eventos de ganancia de función. Este tipo de eventos provocan una reestructuración en las rutas de señalización de la célula cancerígena, haciéndole susceptibles a cambios que en principio, no afectarían a las células normales. La letalidad sintética de dosis, tiene lugar cuando la muerte celular ha sido causada por una mutación de ganancia de función y una inactivación, genética o farmacológica, de un gen o rutas secundarias.
- c. Ganancia de función/Ganancia de función: existen mutaciones de ganancia de función que son excluyentes, por ejemplo, porque activan rutas de señalización opuesta. El estudio de la exclusividad mutua entre genes puede ayudarnos a encontrar dichos genes e inducir muerte celular.



Nature Reviews | Clinical Oncology

Figura 1. Fuente: Drug Therapy: Exploiting synthetic lethality to improve cancer therapy, Diede Brunen and René Bernards, Nature Reviews 2017.

## 2. Basándonos interacciones farmacológicas:

- a. Interacción gen-droga, cuando nos referimos a que la causa de la inhibición secundaria se debe al uso de un fármaco.

- b. Interacción droga-droga: cuando ambas interacciones son debidas a la utilización de fármacos, de forma que podemos hablar de terapias combinadas.

La letalidad sintética no necesariamente tiene que ocurrir como consecuencia de una mutación o alteración molecular. También puede tener lugar como consecuencia del tratamiento con un agente anticancerígeno. Un ejemplo de estas interacciones, es la utilización combinada de inhibidores de BRAF y EGFR.

En pacientes de cáncer de colon, la utilización de terapias directas para la inhibición de BRAF genera resistencias debido a la activación secundaria de la ruta oncogénica de EGFR, capaz de rescatar a las células tumorales. En este caso, la inhibición combinada de ambas rutas mediante fármacos dirigidos, ha demostrado tener efectos beneficiosos para los pacientes (Prahallad A, *et al.*, 2012).

Existe también otro tipo de letalidad sintética que no entra en ninguna de las dos clasificaciones anteriores y que se denomina **letalidad sintética colateral**. Esta definición alude a la utilización de un fármaco primario, con el fin de provocar vulnerabilidades en la célula tumoral, de forma que esta quede sensibilizada a la actuación de un fármaco secundario al cual era previamente inmune. En la clínica, las terapias de segunda o tercera línea, cada vez pueden resultar menos efectivas, pero la identificación de esencialidades colaterales, nos podría llevar al desarrollo de terapias de segunda líneas al menos igual de efectivas que las de primera línea.

### 3.3.2. Retos de la técnica

Como principales desafíos de este acercamiento:

1. Este tipo de relaciones genéticas provocan la muerte de la célula, por lo que su identificación es complicada.
2. Muchas de estas interacciones dependen de las condiciones en las que se encuentren las mutaciones y puede que no sean observadas en todos los perfiles genómicos o en distintas condiciones celulares a las testadas.
3. Además, son interacciones poco frecuentes y se necesita de un gran número de combinaciones de mutantes para poder identificarlas.

### 3.3.3. Detecciones de letalidad sintética en células humanas.

Comparado con el estudio de este tipo de eventos en organismos modelo como la levadura, la interrogación sistemática de pares de letalidad sintética en células humanas, todavía está en sus inicios.

La identificación y caracterización de interacciones de letalidad sintética robustas, es un paso necesario para la exploración de esta posible nueva vía terapéutica (Brunen D, René B, 2017).

Algunas de las fuentes o estrategias que se pueden utilizar son:

1. Análisis de pares de letalidad sintética obtenidos a partir de organismos modelos y estudio de su posible conservación humanos.
2. Predicciones *in silico*.
3. Identificación de dependencias genéticas en células tumorales humanas. Utilizando pruebas de detección masivas a lo largo de paneles de líneas celulares tumorales, la esencialidad génica puede ser mapeada de forma sistemática en los distintos perfiles genómicos. De esta forma, se pueden detectar vulnerabilidades en genes no mutados en cáncer, así como caracterizar su perfil de sensibilidad.

#### 3.3.3.1. Detecciones mediante ARN de interferencia.

El ARN interferente (ARNi) son moléculas de cortas de ARN, de 20 a 30 nucleótidos de largo, capaces de inhibir la expresión o traducción de un gen, mediante la neutralización de las moléculas de ARN mensajero. En concreto, moléculas de ARN de interferencia (ARNi), incluyendo *small interfering RNAs* (siARNs) y *short hairpin RNAs* (shARNs), han generado grandes avances en la identificación sistemática de los efectos fenotípicos (tales como la muerte celular o activación de diversas vías de interés) relacionados con la supresión de genes, permitiendo su detección sistemática (Bernards *et al.*, 2006; Downward, 2004; Westbrook *et al.*, 2005).

La disponibilidad de librerías de ARNi a escala genómica supone una herramienta muy poderosa para la identificación de dependencias específicas de contexto en células cancerígenas (Kittler *et al.*, 2008; Hirsch 2010; Cheung *et al.* 2011). Se busca poder entender mejor la relación entre genes específicos y el comportamiento fenotípico de la célula.

Como ejemplo de los avances que ha supuesto esta técnica, podemos encontrar la investigación de interactores de letalidad sintética para el oncogen KRAS.

KRAS está mutado en el 30% de los tumores humanos y tiene una prevalencia superior al 90% en cáncer de páncreas (Weidle UH *et al.*, 2011). Detecciones masivas de ARN de interferencia han sido utilizados para identificar este tipo de vulnerabilidades.

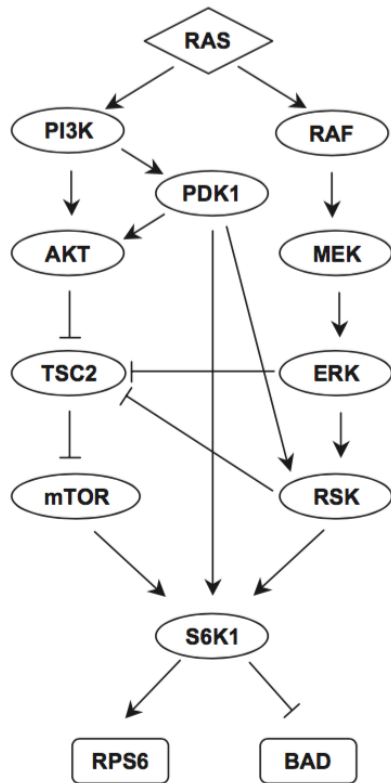


Figura 2. Ruta de interacción entre RAS y STK33. Fuente: Weidle UH *et al.*, 2011.

En este caso, las células que dependen de KRAS son sensibles a la supresión de la serina-treonina kinasa 33 (STK33) independientemente del tejido del que se trate. STK33 promueve la viabilidad de KRAS en células dependientes mediante la supresión de la apoptosis mitocondrial inducida por la proteína ribosomal s6 quinasa 1, la cual inactiva el promotor propapótico BCL2.

Células tumorales con una mutación en NRAS o HRAS, no sufren este tipo de fenómeno, sino que es específico de KRAS. A su vez, STK33 no es un oncogen en sí mismo y la interacción letal entre KRAS y STK33 se restringe a células que son funcionalmente dependientes de KRAS. Para el desarrollo de drogas anticancerígenas basadas en esta interacción, es importante definir el rol de STK33 correctamente, con el fin de entender los posibles efectos tóxicos del agente químico (Scholl C *et al.*, 2009).

### 3.3.3.2. Detecciones mediante CRISPR-Cas9.

CRISPR es una familia de secuencias de ADN de origen bacteriano. Estas secuencias contienen fragmentos de ADN de virus atacantes, que reconocen y se encargan de eliminar, jugando un rol importante en el sistema de defensa bacteriano. Este sistema es la base de la técnica de CRISPR/Cas9, con la cual se consigue cortar genes de forma específica y dirigida.

El análisis de conjuntos de datos mediante esta técnica, tiene el potencial de desenmascarar nuevas interacciones de letalidad sintética. Utilizando CRISPR-Cas9 a lo largo de paneles de líneas celulares, podemos mapear de forma sistemática la esencialidad de genes en distintos contextos, ayudándonos a descubrir genes que no están mutados en

cáncer pero pueden ser esenciales cuando se dan mutaciones específicas de la enfermedad (*non-oncogene addiction*).

### 3.3.3.3. *ARNi vs. CRISPR-Cas9.*

El uso del ARN de interferencia como tecnología para el silenciamiento de genes ha lanzado a la era del *high throughput* los estudios de letalidad sintética.

Pero aunque el avance es claro, la técnica tiene debilidades que deberemos tener en cuenta.

Por un lado, se ha visto existe un solapamiento muy bajo entre estudios independientes en los que se ha usado ARN de interferencia (Bhinder B, Djaballah H, 2013). Esto puede ser debido a uno de los mayores problemas de la técnica, los llamados efectos *off-target*.

Los efectos *off-target* tienen lugar cuando las moléculas de ARNi interfieren en el mecanismo endógeno de silenciamiento de la célula o silencian otros genes, además del gen objetivo. El fenotipo resultante, con todos sus efectos secundarios no deseados, se atribuye al silenciamiento de la diana, dando lugar a un alto número de falsos positivos.

Para intentar contrarrestar estos efectos, han surgido diversas estrategias que implican el desarrollo de nuevos algoritmos bioinformáticos o la utilización de grandes librerías de líneas celulares que ayudarían a normalizar la información obtenida tras cada uno de los silenciamientos.

Otra de las limitaciones de la técnica implica la utilización de librerías de líneas celulares tumorales, modelos que no tienen por qué traducirse a un contexto real y con una alta heterogeneidad, afectando también al bajo solapamiento entre estudios.

Sin embargo, mientras que el ARN de interferencia tiene la desventaja de hacer una inhibición incompleta (*knock-down*), esto puede suponer una ventaja a la larga, pudiéndose establecer de una forma más directa la relación entre los efectos del *knock-down* y los posibles efectos del fármaco al inhibir el gen.

Por su parte, los estudios de silenciamiento de genes mediante CRISPR, han conseguido reducir estos efectos *off-target* debido a un aumento en la especificidad del silenciamiento. Las rutas endógenas relacionadas con el ARN de interferencia no se ven afectadas y por tanto, la identificación de genes esenciales es mucho más directa.

Aún así, experimentos recientes han demostrado que esta técnica es altamente tóxica para las células tratadas, pudiendo afectar el fenotipo resultante (Munoz DM *et al.*, 2016).

Es por ello que podemos decir que existen puntos a favor y en contra de ambas técnicas y que la información extraída a través de ambas estrategias es igualmente valiosa.

#### **3.3.4. *De la detección de interacciones de letalidad sintética a la práctica clínica.***

Como una de las grandes ventajas que suponen este acercamiento, los pares de letalidad sintética son específicos de mutaciones presentes exclusivamente en células cancerígenas. Es por ello que permite un mayor control de la dosis de la terapia aplicada, suponiendo una disminución en los efectos secundarios derivados del tratamiento.

Además, al actuar sobre dianas secundarias, muchas inexploradas en el tratamiento del cáncer, esta estrategia supone un aumento de la ventana terapéutica disponible.

También es importante señalar que la estrategia puede ser aplicada a cualquier tipo de mutación, incluyendo genes supresores de tumores y otros genes *undruggable*.

Claramente, hay que determinar parámetros adicionales como efecto del quimioterapéutico, los posibles mecanismos de resistencia que pueden tener lugar y entender los mecanismos biológicos que explican esa interacción, antes de considerarlos para su aplicación. Ahora es el momento de re-evaluar cómo priorizar estos candidatos para su testeo terapéutico.

Pese a los avances que están teniendo lugar, todavía existen muy pocas drogas aprobadas para su uso terapéutico basado en el concepto de letalidad sintética. Aun así, se espera que las detecciones a gran escala utilizando CRISPR-Cas9 acelerarán el descubrimiento de nuevas terapias y dianas terapéuticas, así como combinaciones de tratamientos.

Estas interacciones en muchas ocasiones, dependen de contexto y por tanto, el éxito de los tratamientos de letalidad sintética en ensayos clínicos, dependerá de la selección de los pacientes según el contexto molecular y características de su tumor.

Además, muchos estudios han demostrado que la adquisición de resistencias en terapias dirigidas suele ser bastante rápida. Las detecciones de SL son ideales para mapear estos mecanismos de resistencia intrínsecos y diseñar combinaciones de drogas que pueden evitarlos.

## 4. Materiales y métodos

### 4.1. Conjuntos de datos.

A continuación se describen brevemente, las bases de datos utilizadas para el desarrollo de este proyecto.

#### 4.1.1. *SynLethDB*

Se trata de una base de datos dedicada a la recolección de pares de SL identificados tanto en humano, como en cuatro organismos modelo.

SynLethDB integra diferentes metodologías de detección y en el caso de los pares provenientes de estudios en humano, realiza un análisis estadístico de la sensibilidad, de las distintas líneas celulares tumorales, al tratamiento mediante fármacos disponibles en: DrugBank database, STITCH (una base de datos que contiene información de interacciones droga-proteína) y KIBA (base de datos de perfiles de inhibición droga-quinasa).

También ofrece la posibilidad de buscar información de ortología, integrando la información de los demás organismos modelos.

Las fuentes de información integradas en SynLethDB son las siguientes:

1. Información curada manualmente a partir de artículos de investigación.
2. Pares identificados a partir de experimentos de secuenciación masiva: datos de experimentos combinatorios entre RNAi y testeos de sensibilidad drogas o screenings de shRNA provenientes del proyecto DECIPHER (<http://www.decipherproject.net/>).
3. Interacciones genéticas anotadas como “SL pair” en la base de datos de BioGRID (<https://thebiogrid.org/>).
4. Pares de genes anotados como “SL” en la base de datos de GenomeRNAi (<http://www.genomernai.org/>), un proyecto que recopila fenotipos sacados de estudios de RNAi en *Drosophila* y *Homo sapiens*.
5. Pares de letalidad sintética predichos por DAISY (Ryan CJ *et al.*, 2014).
6. Resultados de la búsqueda de los términos “synthetic lethal” y “synthetic lethality” mediante técnicas de Text mining.

Una vez integrada toda esta información, SynlethDB otorga una puntuación a cada par de letalidad sintética, con el fin de establecer cuales recogen más evidencias y por tanto, qué interacciones son más fiables. Estas puntuaciones no han sido exploradas en conjuntos de datos externos, sino que se basan puramente en la información que ya estaba disponible.

Básicamente, según los criterios que han establecido para otorgar las distintas puntuaciones:

1. Las evidencias experimentales contribuyen de forma más significativa a la puntuación final que los algoritmos de predicción (como DAISY) o text mining.
2. Se otorga peso diferente a los estudios experimentales según las técnicas que se hayan utilizado para obtener dichas evidencias, considerando menos fiables los screenings masivos.
3. Cuantas más evidencias existen, más puntuación tendrán los pares y por tanto, se consideran más fiables.

Estos pesos han sido integrados y normalizados con el fin de obtener una puntuación final que oscila entre 0 y 1, indicando esta última una mayor confianza.

#### **4.1.2. *PanDrugs***

PanDrugs (<http://www.pandrugs.org/#/>) es un método para la selección de terapias a partir del análisis de alteraciones moleculares presentes en cáncer.

La aproximación de PanDrugs ha permitido la implementación de una base de datos dedicada a la identificación de alteraciones moleculares accionables, ofreciendo aquellas opciones terapéuticas que tengan dichas alteraciones como diana y priorizando los resultados en base a una serie de criterios entre los que se incluyen: información biológica, información sobre la relevancia clínica de los genes de interés, así como datos de susceptibilidad a la aplicación de las diversas soluciones terapéuticas disponibles.

PanDrugs también incluye información sobre la familia de la droga sugerida y su estado de aprobación (ya sea una terapia aprobada, en ensayo clínico, en estado experimental o retirada). Asimismo, la herramienta permite encontrar información sobre el área patológica de la droga ofrecida, definición de si se trata de un target directo o un biomarcador de respuesta a tratamiento y también se incluye entre otras cosas, información de resistencia o sensibilidad a droga y el tipo de alteración molecular.

Con esta información se busca ayudar en la toma de decisiones a la hora de estudiar distintas opciones terapéuticas.

PanDrugs integra datos de 18 fuentes distintas, con los distintos niveles de información de los que ya hemos hablado, e incluye datos provenientes de estudios experimentales en líneas celulares de cáncer. Actualmente contiene ~50,000 asociaciones droga-diana para ~6,000 genes y ~11,000 compuestos.



Entre estas bases de datos, podemos encontrar:

1. DGIdb: catálogo de información sobre gene druggability.
2. Monoclonal Antibodies
3. TARGET
4. Cancer Therapeutics Response Portal (CTRP)
5. Genomics of Drug Sensitivity in Cancer (GDSC)

#### 4.1.3. *Project Achilles*

Project Achilles es un proyecto desarrollado por el Broad Institute (Massachusetts, E.E.U.U.), destinado a generar un catálogo de esencialidad génica para cientos de líneas celulares relacionadas con cáncer.

Para llevar a cabo esta misión, el proyecto utiliza librerías shARN y CRISPR/Cas9 con el fin de suprimir de forma estable de cada uno de los genes del genoma completo.

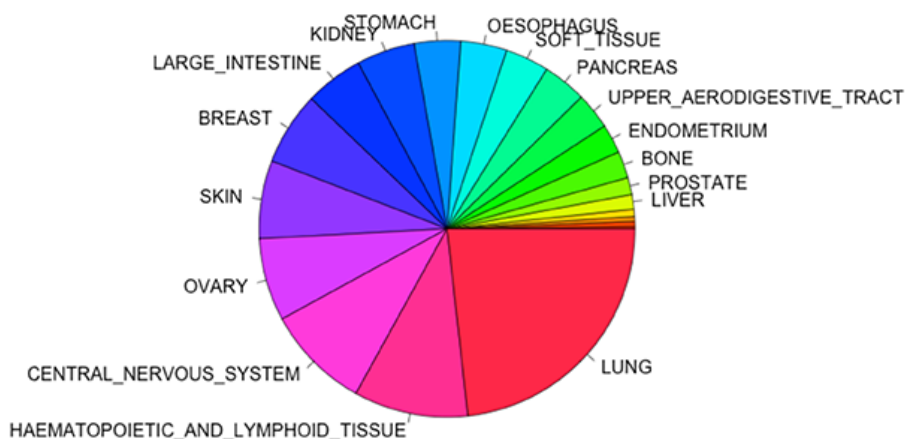


Figura 3. Origen tumoral de las muestras incluidas en la versión 2.20.2 de Achilles. Fuente: <https://portals.broadinstitute.org/achilles/about>

Actualmente, la base de datos dispone de información de silenciamientos, para más de 10,000 genes, mediante RNAi para 501 líneas celulares y mediante CRISPR-Cas9, para 342 líneas (Meyers RH *et al.*, 2017).

#### 4.1.4. *Cancer Cell Line Encyclopedia*

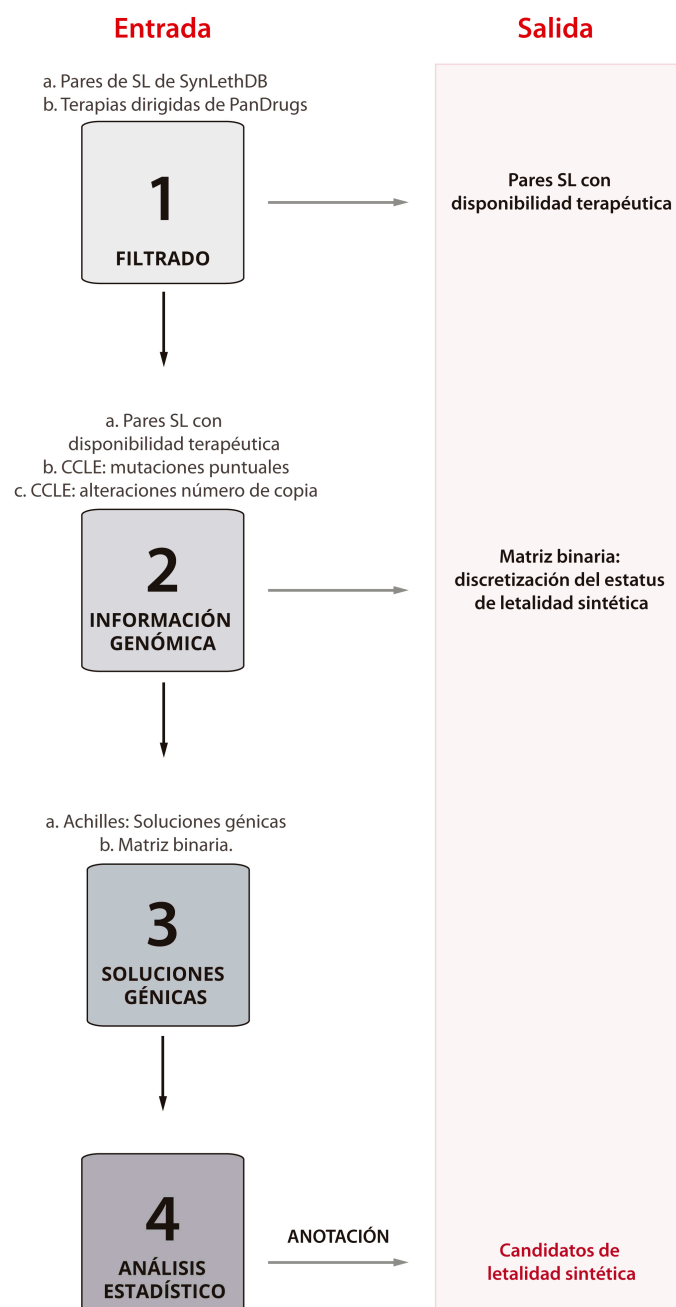
El proyecto de la Cancer Cell Line Encyclopedia (CCLE) consiste en la caracterización genética de un panel de más de 1000 líneas celulares tumorales de origen humano.

Para el análisis de los datos provenientes de Project Achilles, ha sido utilizada la información tanto de mutaciones puntuales como de alteración en número de copia, disponible para las 501 líneas celulares del estudio en la Cancer Cell Line Encyclopedia.

#### 4.2. Integración de datos.

Toda la información en la que se ha basado este estudio, ha sido descargada de forma manual. Para el parseado y aplicación de filtros se ha utilizado *R programming*.

Los pasos seguidos han sido los siguientes:



#### 4.2.1. Tratamiento de los datos.

Nuestra estrategia es la siguiente: queremos explorar los pares de letalidad sintética anotados en SynLethDB y con un tratamiento disponible, utilizando la información de soluciones génicas obtenida del *Project Achilles*.

Para comprobar si un par de letalidad sintética es efectivo en una línea celular determinada, necesitamos anotar el estado mutacional de los dos genes que lo conforman. Para ello, vamos a usar la información de alteraciones en número de copia y mutaciones puntuales disponible para esa línea en la base de datos del CCLE.

En SynLethDB, cada par de letalidad sintética está anotado como “gen A” y “gen B”, existiendo una direccionalidad que hemos decidido mantener para este estudio.

Básicamente, necesitamos que el “gen A” esté mutado, de forma que podamos comprobar cuánto se ve afectada la viabilidad celular cuando se silencia el “gen B” mediante ARN de interferencia.

Priorizaremos aquellos casos en los que la mutación de ambos genes provoque una disminución en la viabilidad de la célula, de esta forma, pretendemos conseguir un curado de las relaciones de letalidad sintética de mayor interés biológico.

GeneASymbol	GeneAId	GeneBSymbol	GeneBId	PubmedID	Evidence	Type Species	Disease	Score
BRCA1	672	PARP1	142	101337;21777194;225621	ethality;Decipher;Text Mli	SL	lan cancer;sporadic cancer	0,99
BIRC5	332	HRAS	3265	12833149	Synlethality	SL	Human cancer	0,9
BRCA1	672	PARP2	10038	21487248;3965078	Synlethality;Decipher	SL	an cancer;A549,PC3,MDA	0,873
RB1	5925	TOP2A	7153	11313881	Synlethality	SL	Human cancer	0,855
BRCA2	675	PARP2	10038	15829967	Synlethality	SL	Human breast cancer	0,855
ECT2	1894	RB1	5925	16862181	Synlethality	SL	Human cancer	0,855
KRAS	3845	VDAC1	7416	17568748	Synlethality	SL	Human lung cancer	0,855
MSH2	4436	POLB	5423	20227038	Synlethality	SL	Human colorectal cancer	0,855
EGFR	1956	PRKACB	5567	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	PRKCD	5580	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	PRKCE	5581	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	PRKCZ	5590	20858866	Synlethality	SL	Human cancer	0,855
ABL1	25	EGFR	1956	20858866	Synlethality	SL	Human cancer	0,855
CALM1	801	EGFR	1956	20858866	Synlethality	SL	Human cancer	0,855
CD59	966	EGFR	1956	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	FGFR2	2263	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	PIK3R1	5295	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	PIK3R2	5296	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	RET	5979	20858866	Synlethality	SL	Human cancer	0,855
EGFR	1956	SMAD2	4087	20858866	Synlethality	SL	Human cancer	0,855
DUSP7	1849	EGFR	1956	20858866	Synlethality	SL	Human cancer	0,855
BCAR1	9564	EGFR	1956	20858866	Synlethality	SL	Human cancer	0,855
HRAS	3265	PRKCD	5580	21031151	Synlethality	SL	Human prostate cancers	0,855
KRAS	3845	MAPK14	1432	21306997	Synlethality	SL	non-small cell lung cancer	0,855

Figura 4. Esta tabla recoge los datos disponibles para muestras de humano en SynLethDB. La tabla ha sido ligeramente modificada para mejorar la comprensión de la información que contiene. Entre los campos de mayor interés podemos encontrar: GeneASymbol y GeneBSymbol, que nos indican el nombre de los genes que forman el par de letalidad sintética; Evidence, que nos informa de las fuentes de información que recogen dicho par; y Score, que nos indica la puntuación otorgada para cada uno de los pares incluidos en la base de datos.

#### **4.2.2. Paso 1: Cruce entre Pandrugs y SynLethDB.**

En primer lugar, se obtuvieron los fármacos provenientes de terapias dirigidas de la última versión de la base de datos de Pandrugs. A partir de este listado, hemos cruzado los “genes B” con las dianas de todas las terapias seleccionadas. Así, el resultado final es un listado de pares de letalidad sintética y posibles terapias efectivas, filtrando todos aquellos candidatos que no tuvieran un posible tratamiento.

#### **4.2.3. Paso 2: Filtrado de los pares de SL de interés mediante el uso de información genómica.**

A partir del listado de pares de SL con disponibilidad terapéutica, se ha realizado un filtrado adicional en base a los datos de alteraciones en número de copia y mutaciones puntuales para cada una de las líneas celulares estudiadas. Esta información fue obtenida del portal del CCLE.

El filtrado ha consistido en el estudio mutacional de cada uno de los componentes del par de letalidad sintética analizado. El resultado es un listado de pares en los que se cumplen las siguientes condiciones: que el “gen A” esté mutado y que el “gen B” no lo esté.

##### **4.2.3.1. Para la información de mutaciones puntuales.**

Hemos estudiado exclusivamente aquellas mutaciones puntuales que no fueran de tipo silencioso. El hecho de tener algún tipo de impacto ya fuera: de cambio de sentido, sin sentido, o con desplazamiento de marco de lectura, es suficiente para determinar que el gen de interés está mutado y proseguir con los siguientes pasos del análisis.

##### **4.2.3.2. Para el análisis de CNVs.**

Hemos discretizado los datos de número de copia provenientes del CCLE siguiendo estos criterios:

1. Consideraremos deleciones profundas todas aquellas con un valor de número de copia inferior a 0.1
2. Consideraremos deleciones, todos los valores que oscilan entre 0.1 y 0.5.
3. Consideraremos deleciones hemigotas (de una sola de las copias) todos los valores que oscilan entre 0.5 y 1.
4. Los valores entre 1 y 2, han sido considerados como valores normales, es decir no se ha producido ninguna deleción o amplificación.
5. Entre 2 y 8 copias, consideramos que el gen ha sido amplificado.
6. Consideramos grandes amplificaciones, aquellos genes que tienen más de 8 copias.

Sólo aquellos casos en los que se haya producido una: delección profunda, delección o gran amplificación, serán tenidos en cuenta para los siguientes pasos.

#### 4.2.3.3. **Consenso.**

Finalmente, hemos establecido un consenso entre las anotaciones de mutaciones puntuales y de número de copia, de forma que el resultado final son pares de letalidad sintética con una mutación de cualquier origen, en el gen A y ninguna en el gen B.

Mediante esta matriz final, sabremos en qué líneas celulares se encuentran estos pares y en cuáles no.

#### 4.2.4. **Paso 3: Análisis de viabilidad celular.**

Para analizar el efecto de las interacciones de SL seleccionadas, se ha integrado la información de esencialidad génica obtenida del Project Achilles.

La corrección de efectos *off-target* derivados de los experimentos del Project Achilles que utilizan librerías de shRNA se ha realizado utilizando el método computacional DEMETER, desarrollado por Aviad Tsherniak (Tsherniak A *et al.*, 2017). El resultado de la aplicación de este algoritmo son las llamadas **soluciones génicas** (en inglés, *gene solutions*), valores que representan las consecuencias fenotípicas que suponen eliminar el gen en la muestra de interés, normalizado por las consecuencias en el resto de líneas celulares. Esta medida básicamente nos informa del grado en el que se ha visto afectada cada línea celular, siendo los valores más bajos los que representan un mayor efecto en la viabilidad de la línea estudiada. Todo aquel valor inferior a -2 se considera que provoca la muerte celular (Tsherniak A *et al.*, 2017).

En primer lugar, en base a los filtros aplicados en el paso anterior y para cada uno de los pares interactores de estudio, hemos seleccionado cuáles deberían tener consecuencias de tipo SL y cuáles no, generando dos grupos. Posteriormente, hemos analizado la distribución de los grupos creados según los datos de esencialidad.

#### 4.2.5. **Paso 4: Análisis estadístico.**

Mediante la prueba U de Mann-Whitney, una prueba no paramétrica utilizada para analizar grupos independientes, hemos examinado si existen diferencias significativas en los niveles de esencialidad de los pares seleccionados como SL y el resto de muestras.

En otras palabras, para cada par de genes denominados A y B, la prueba U de Mann-Whitney se puede llevar a cabo con el fin de analizar si la inhibición del gen B mediante shRNA, lleva a una disminución en el score de esencialidad génica en aquellas muestras en las que el gen A está mutado (ya sea inactivado o sobre-activado), comparándolo con el resto de las muestras. Este tipo de test ha sido utilizado para analizar relaciones de letalidad sintética en estudios como DAISY (Ryan CJ *et al.*, 2014) o CancerGD (Bridgett S *et al.*, 2017).

#### **4.2.6. Paso 5: Caracterización y anotación.**

Con el fin de conocer la funcionalidad de los pares de letalidad sintética en un contexto de cáncer, los genes A de nuestra selección fueron anotados utilizando la librería de *Cancer Gene Census* (Futreal PA *et al.*, 2004) y *ONGene* (Liu Y *et al.*, 2017).

La librería de *Cancer Gene Census*, consiste en un listado de genes “driver” curado manualmente, para los cuales se existe información sustancial y publicada sobre su mecanismo de acción en diversos tipos de cánceres humanos, ya sea como oncogenes o como genes supresores de tumores. Sólo aquellos para los cuales hay evidencias de al menos dos estudios independientes, están incluidos en el listado.

Por otro lado, *ONGene* es una base de datos curada manualmente a partir de la literatura, de oncogenes humanos.

## 5. Resultados

### 5.1. Resultados derivados del pipeline de filtrado

#### 5.1.1. Genes con terapias directas seleccionados

Para este estudio hemos partido de las interacciones de letalidad sintética en humano, disponibles en SynLethDB. En total, se tratan de 16,916 interacciones, generadas a partir de 3,316 genes A y 3,241 genes B.

Mediante el primer paso del filtrado, hemos seleccionado aquellas interacciones que son susceptibles a la modulación mediante fármacos.

Obteniendo un total de 5,890 relaciones, para las que existe una droga disponible capaz de atacar de forma directa al gen B.

Estas relaciones equivalen a un 34% del total y surgen de la interacción entre 2,207 genes A y 568 genes B diferentes.

De los 2,207 genes A seleccionados, un 78,2% (1726 en total) eran genes previamente *undruggable* mediante terapias directas, es decir, no estaban presentes en la base de datos de PanDrugs y por tanto, no tenían una terapia disponible. De estos, 1302 genes son genes previamente *undruggable* tanto por terapias directas como utilizando el acercamiento de rutas desarrollado para la base de datos PanDrugs.

Es decir, añadiendo estas nuevas interacciones, estaríamos enriqueciendo con más de 1300 genes y sus correspondientes opciones terapéuticas, la base de datos de PanDrugs.

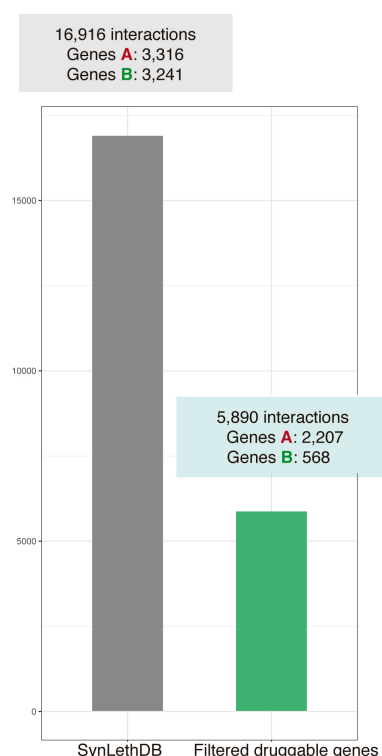


Figura 4. En este histograma, podemos ver las interacciones SL con las que iniciamos el análisis (anotadas como SynLethDB) y las interacciones que conforman nuestra selección de pares de letalidad sintética con opción u opciones terapéuticas disponibles (anotadas como Filtered druggable genes). Esta figura ha sido generada mediante el paquete ggplot2, a través del software de análisis R.

### 5.1.2. Fuentes de origen de los datos

Todos estos genes, provienen de diversas fuentes, que también hemos analizado con el fin de entender mejor la naturaleza de los datos.

Como ya hemos comentado, SynLethDB otorga un peso específico a cada una de las interacciones anotadas, siendo las interacciones demostradas experimentalmente las que más peso tienen; y las predicciones bioinformáticas, junto con la información determinada mediante herramientas de minería de datos (*text mining*), las que menos.

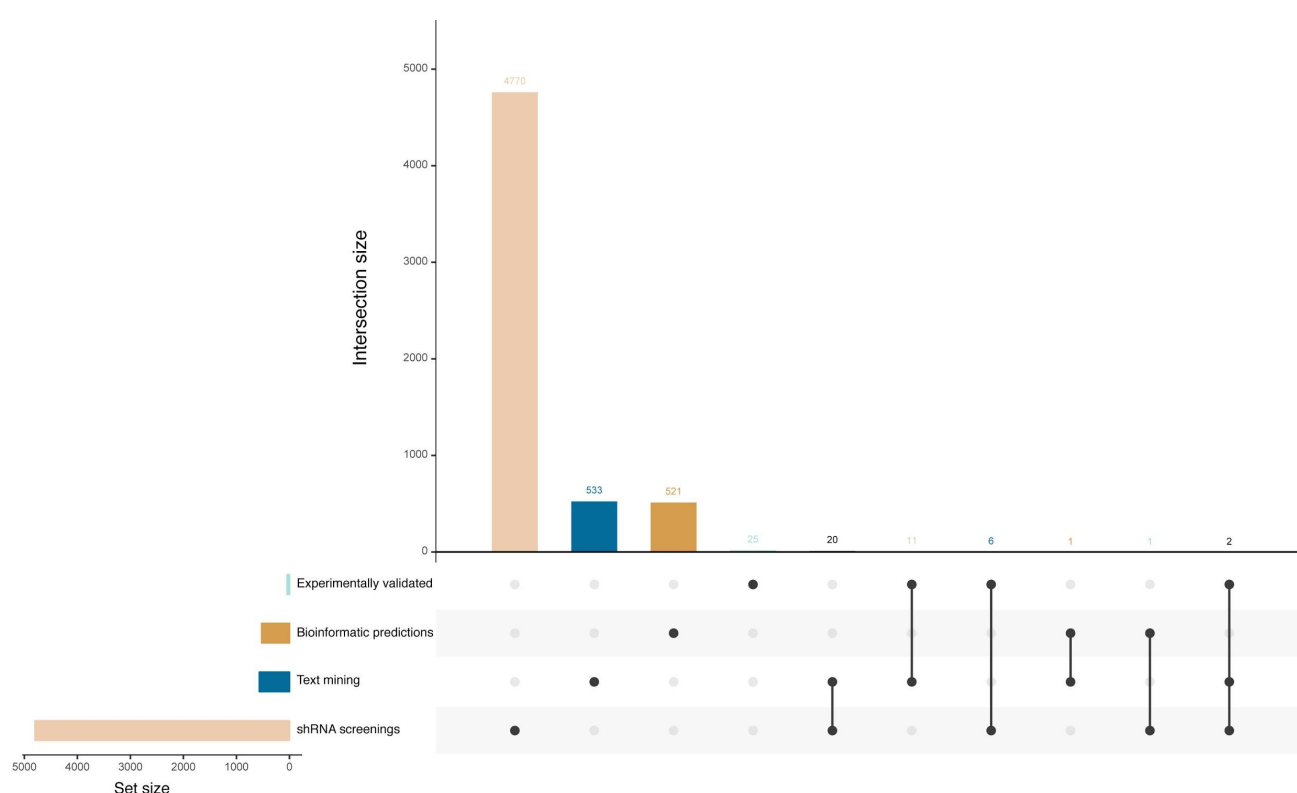


Figura 6. Análisis de las fuentes de origen (Evidences) de los datos filtrados.

La categoría *Experimentally validated* se refiere a aquellos pares que han sido validados experimentalmente; *Bioinformatic predictions*, a los pares de SL obtenidos a través de predicciones mediante el algoritmo bioinformático Daisy; *shRNA screenings* contiene en su categoría tanto pares obtenidos mediante el proyecto GenomeRNAi, como mediante el proyecto DECIPHER. *Text mining* incluye aquellos pares localizados tras la aplicación de filtros de minería de datos utilizando el término “synthetic lethality” o “synthetic lethal”. La figura ha sido generada mediante el paquete UpsetR.

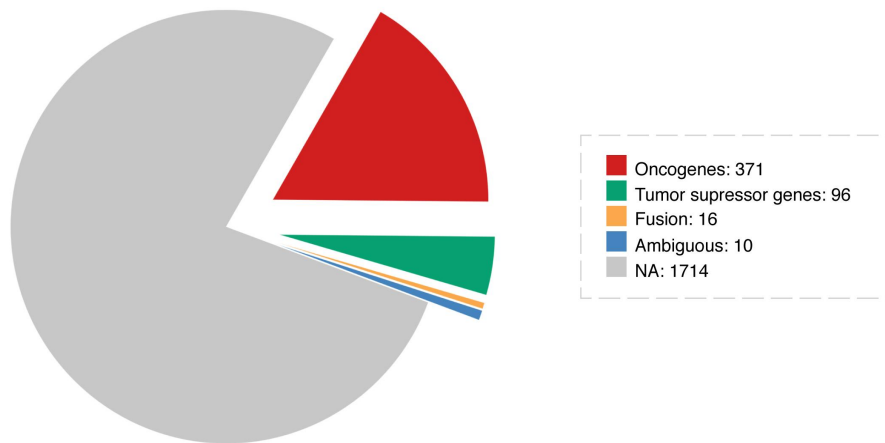
En esta figura, podemos ver que la mayoría de interacciones *druggable* seleccionadas, tienen su origen en proyectos de silenciamiento de genes, entre los que se incluyen las interacciones provenientes de GenomeRNAi y de DECIPHER. Existiendo sólo 44 interacciones provenientes de estudios validados experimentalmente.



## 5.2. Anotación

De los 2,207 genes A seleccionados, nos interesaba conocer un poco mejor su papel en cáncer. Esta información nos ayudará en el análisis de las distintas estrategias terapéuticas propuestas.

Una vez anotados mediante el uso de Cancer Genome Census y ONGene, 493 de ellos resultaron genes con funciones causales en cáncer conocidas: 96 son genes que actúan como supresores tumorales y 371, como oncogenes.



*Figura 7. Este diagrama muestra todos los genes A de nuestro análisis y su anotación (en el caso de tenerla) de su funcionalidad en cáncer. El diagrama ha sido generado mediante la herramienta ggplot2.*

## 5.3. Tabla de resultados

Como resultado final y tras la aplicación de nuestros filtros, hemos obtenido una tabla anotada en la que podemos encontrar la siguiente información para todos los pares susceptibles a la modulación mediante fármacos:

- Para el gen A: nombre, líneas celulares en las que está mutado, mutaciones puntuales y de número de copia en dichas líneas, anotación de su implicación en cáncer.
- Para el gen B: nombre, líneas celulares en las que está mutado, mutaciones puntuales y de número de copia en dichas líneas.
- Para la interacción: información derivada de la base de datos SynlethDB (fuente y score) y p-valor obtenido para la asociación.

## 5.4. Casos de estudio

### 5.4.1. Caso 1: BRCA1 - PARP1

Entre nuestros resultados, podemos encontrar ejemplos muy interesantes, como es el de la interacción entre BRCA1 y PARP1 (p-valor: 0.0441).

BRCA1, junto a BRCA2, son genes supresores de tumores cuyo producto proteico cumple un papel fundamental en la reparación del ADN mediante recombinación homóloga, ayudando a la regulación del ciclo celular.

Cuando BRCA está mutado, el ADN no se repara de forma correcta, por lo que las células empiezan a acumular alteraciones adicionales que pueden llevar al desarrollo de un tumor.

Se estima que la mutación, tanto de BRCA1 como de BRCA2, está implicada en el desarrollo de un 5-10% del total de cánceres de mama (Campeau PM *et al.*, 2008) y 15% de ovario (Pal T *et al.*, 2005). Estos

cánceres son especialmente agresivos debido a la falta de terapias dirigidas a atacar esta pérdida funcional.

Por su parte, la poli ADP-ribosa polimerasa 1 (PARP1), se encarga de la reparación del ADN en aquellas células donde BRCA está mutado. Es decir, las células que tienen mutaciones en BRCA, dependen de PARP para seguir siendo viables y por ello, el desarrollo de fármacos inhibidores de PARP, ha tenido un papel fundamental en el tratamiento de este tipo de tumores.

La inhibición de PARP provoca una reparación del ADN ineficiente, que lleva a las células tumorales a la apoptosis. Las células normales por su parte, que no deberían tener mutado este supresor tumoral, podrían seguir desarrollando sus funciones de reparación y por tanto, sobrevivir a la terapia.

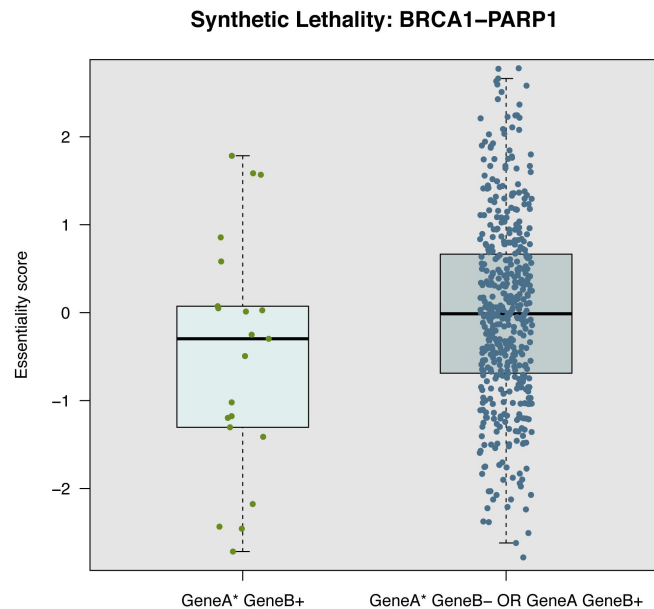


Figura 8. En este boxplot podemos ver todos aquellos casos clasificados como SL a la izquierda (Gene A\* - GeneB+, nos indican que el gen A está mutado, mientras que el B no lo está) y el resto de muestras, en el boxplot de la derecha.

La prueba U de Mann-Whitney ha sido aplicada a estos dos grupos con el fin de testar si existían diferencias significativas entre la media de ambas clasificaciones. El p-valor obtenido para esta interacción es 0.0441.

No sólo la interacción entre BRCA1 y la poli ADP-ribosa polimerasa 1 (PARP1) es una de las mejor conocidas al hablar de letalidad sintética, sino que el uso de inhibidores de PARP en pacientes con mutaciones en BRCA, ha sido la primera estrategia basada en letalidad sintética y aprobada para su uso en cáncer.

Esta terapia supone que células con mutaciones de pérdida de función en supresores tumorales, puedan ser atacadas mediante el uso de fármacos dirigidos a inhibir las rutas secundarias.

#### 5.4.2. Caso 2: ERBB2 - PARP1

ERBB2 es un oncogen amplificado en un 15-20% de los casos de cáncer de mama y cáncer gastrointestinal (Yersal O, Barutca S, 2014). El producto génico de la sobreexpresión o amplificación de ERBB2 suele estar asociado con un mal pronóstico, provocado por una rápida tasa de crecimiento y el desarrollo de metástasis, generalmente en el cerebro en un 50% de los pacientes con esta alteración (Lin NU, 2013).

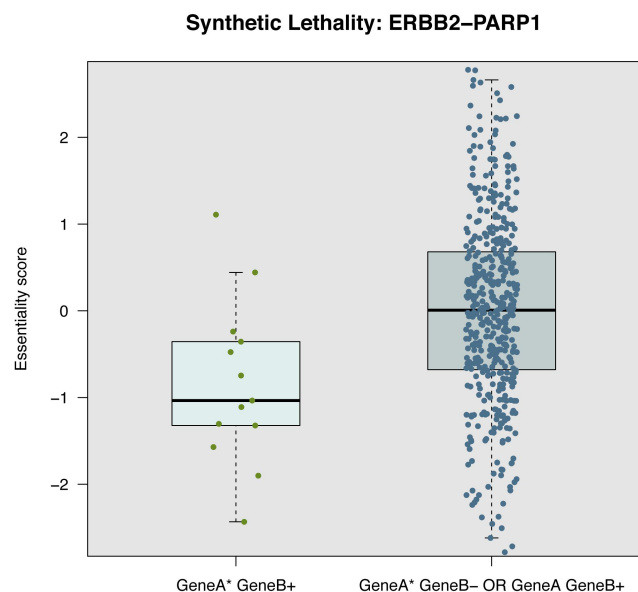
La tasa de supervivencia de los pacientes de cáncer de mama y cáncer gastrointestinal se ha conseguido mejorar mediante la aplicación de *Trastuzumab*, una terapia dirigida a la inhibición de

ERBB2.

Se sabe que la sobreexpresión de ERBB2 en algunos tipos celulares puede aumentar el estrés en la replicación del ADN, un contexto en el que la inhibición

de los mecanismos de reparación del ADN llevados a cabo por PARP puede surgir como una estrategia para provocar la muerte de las células tumorales con dicha mutación.

De hecho, existen ejemplos donde esta interacción de letalidad sintética ya ha sido demostrada (Kawahara N *et al.*, 2017; Nowsheen S *et al.*, 2012) y experimentos en los que



*Figura 9. En este boxplot podemos ver todos aquellos casos clasificados como SL a la izquierda (Gene A\* - GeneB+, nos indican que el gen A está mutado, mientras que el B no lo está) y el resto de muestras, en el boxplot de la derecha.*

La prueba U de Mann-Whitney ha sido aplicada a estos dos grupos con el fin de testar si existían diferencias significativas entre la media de ambas clasificaciones. El p-valor obtenido para esta interacción es 0.00235.

también se ha demostrado que la utilización de inhibidores de PARP aumenta los efectos antitumorales de *trastuzumab* en pacientes con sobre-expresión de ERBB2, abriendo una puerta prometedora para el desarrollo de estrategias para el tratamiento de pacientes con cáncer de mama y gastrointestinal (García-Parra J *et al.*, 2014).

#### 5.4.3. Caso 3: KRAS - UGCG

KRAS es un proto-oncogen cuya proteína desarrolla funciones esenciales en la señalización celular. Su mutación es necesaria para el desarrollo de diversos tipos de cáncer. Está mutado en un 30% de los tumores y tiene una prevalencia del 90% en cáncer de páncreas (Munoz DM *et al.*, 2016)

Hasta la fecha, no existe una terapia que permita la inhibición farmacológica de KRAS. Es por ello que la búsqueda de estrategias basadas en letalidad sintética para la inhibición de este gen es tan importante (Cox DA *et al.*, 2014).

Por su parte, la UDP-glucosa ceramida glucosiltransferasa (UGCG) es un gen cuyo producto proteico está

relacionado con la producción de esfingolípidos. Concretamente, la UGCG, se encarga del primer paso de glicosilación en la biosíntesis de glicosfingolípidos.

Los esfingolípidos desempeñan un papel fundamental en la modulación de la fluidez de la membrana celular y en señalización, teniendo un papel importante procesos relacionados con proliferación y apoptosis (Pitson SM, 2011).

UGCG cataliza la glucosilación de la ceramida, encargada de desempeñar funciones de señalización como supresor tumoral. Este esfingolípido es transformado mediante un proceso enzimático de dos etapas, en esfingosina 1-fosfato (S1P), un esfingolípido relacionado con funciones pro-angiogénicas y de supervivencia. Es el balance entre estos

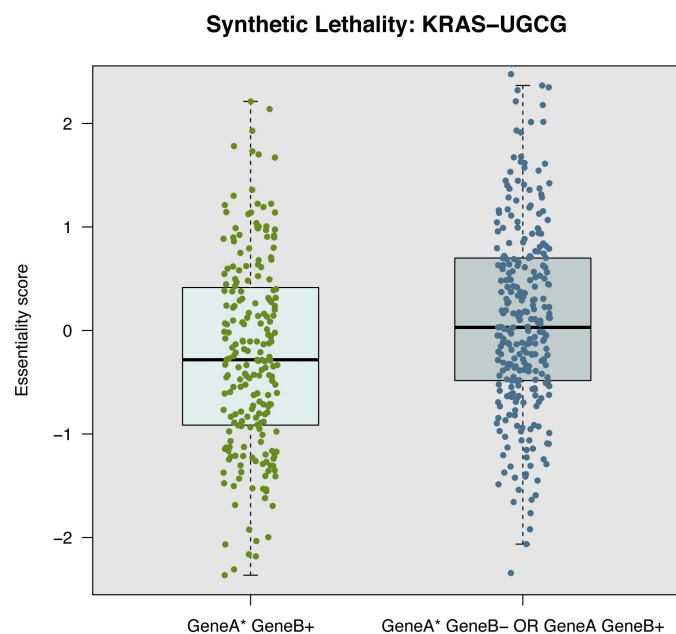


Figura 10. En este boxplot podemos ver todos aquellos casos clasificados como SL a la izquierda (Gene A\* - GeneB+, nos indican que el gen A está mutado, mientras que el B no lo está) y el resto de muestras, en el boxplot de la derecha.

La prueba U de Mann-Whitney ha sido aplicada a estos dos grupos con el fin de testar si existían diferencias significativas entre la media de ambas clasificaciones. El p-valor obtenido para esta interacción es 3.375E-05.

dos metabolitos, el encargado de controlar la supervivencia o apoptosis de las células cancerígenas (Hannun YA *et al.*, 2008; Henry B *et al.*, 2013).

La S1P se suele producir intracelularmente y es exportada por diversos mecanismos (dependientes o independientes de ATP) al exterior de la célula (Kobayashi *et al.*, 2006). Una vez en el exterior de la célula, S1P se une a proteínas G iniciando una cascada de reacciones entre las que se incluyen la maduración vascular, angiogénesis, inflamación, regulación del tráfico de linfocitos en sangre y órganos linfáticos (Rosen H, Goetzl EJ, 2005).

En los últimos años se ha empezado a establecer el rol de S1P en el desarrollo de diversos tipos de cáncer y el papel fundamental que desarrolla durante la tumorigénesis, contribuyendo a la progresión y proliferación del tumor (Xu *et al.*, 2001, Shida *et al.*, 2008).

Además, se sabe que S1P induce a la autofagia, de forma que actúa como protector ante posibles agentes citotóxicos (Lavieu *et al.*, 2006; Chang *et al.*, 2009), pudiendo tener un rol en el desarrollo de resistencias a fármacos (Sukocheva *et al.*, 2009; Pitson, 2011).

La activación de UGCG por su parte, está relacionada con la adquisición de resistencia a quimioterapia en células tumorales, debido a la acumulación de glucosilceramida se acumula en células resistentes a multi-fármacos *in vitro* (Nicholson KM *et al.*, 1999, Lucci A *et al.*, 1998).

No hay mucha información sobre el papel de UGCG *in vivo*, pero existen experimentos de detección por los cuales se ha visto que la inhibición o silenciamiento del gen, restaura la sensibilidad a drogas en células resistentes de leucemia crónica mieloide (Huang WC *et al.*, 2011), mediante un aumento de la concentración celular de ceramidas.

También, existen evidencias de que la inhibición de BCL2, un proto-oncogen implicado en el bloqueo de la apoptosis, y UGCG han resultado en una muerte celular sinérgica (Casson *et al.*, 2013). Por lo que esta interacción de letalidad sintética descrita, resulta de especial interés. Mediante dicha estrategia, se pretende inducir la muerte celular en células con mutaciones oncogénicas en KRAS, haciendo posible la aplicación de una terapia dirigida en pacientes.

### **Reposicionamiento fármacos**

Si estudiamos este caso un poco más a fondo y buscando las opciones terapéuticas que ofrece PanDrugs para inhibir UGCG, el fármaco elegido por la herramienta como primera opción es Miglustat.

Miglustat es un fármaco aprobado para el tratamiento de la enfermedad de Gaucher tipo 1, una enfermedad rara producida por la acumulación de glucosilceramida, en las paredes de células y órganos. Este fármaco fue aprobado en Europa en 2002 y por la FDA en 2003.

La enfermedad de Gaucher de tipo I, es un desorden recesivo autosómico. Los afectados por esta enfermedad tienen un defecto en la enzima glucocerebrosidasa, encargada de convertir glucocerebrósido (o glucosilceramida) en ceramida y glucosa. Cuando esta enzima no funciona, moléculas de glucocerebrósido se acumulan, causando desórdenes en el tamaño de hígado y bazo del enfermo, así como desórdenes en huesos, sangre y médula ósea. El papel de Miglustat consiste en la prevención de la formación de glucoceramidasa.

La capacidad antitumoral de Miglustat también ha sido evaluada en modelos celulares de cáncer con cierto éxito. Pero los efectos off-target de la droga, que entre otras cosas, afectan a la regulación del sistema inmune, ponen en duda su aplicabilidad (Barth *et al.*, 2013).

Sí que se ha demostrado que el uso de análogos de la glucosilceramida, tienen una eficacia significativa en modelos in vitro e in vivo de la eficacia antitumoral en modelos in vitro e in vivo de resistentes a terapia (Liu J *et al.*, 2013). Esta eficacia deriva de la capacidad que estos análogos tiene para bloquear la glucosilceramida sintasa y la P-glicoproteína (encargada de la expulsión dependiente de ATP de esfingolípidos a través de la membrana). Con este mismo propósito, podría ser interesante tener en cuenta el efecto del quimioterapéutico *tamoxifen*, un inhibidor también de la P-glicoproteína y de la glucosilación de la ceramida.

En resumen, la exploración de estas interacciones de letalidad sintética nos pueden llevar a la obtención de nuevas evidencias y dianas terapéuticas, ayudando a profundizar en los procesos y mecanismos biológicos que se encuentran detrás de la enfermedad, así como mejorando nuestro entendimiento en el estudio de las opciones terapéuticas disponibles.

## 6. Discusión

La identificación de nuevas dianas terapéuticas, es uno de los mayores retos en el campo de la medicina personalizada.

En los últimos años, ha surgido una estrategia prometedora basada en el concepto de letalidad sintética, que puede suponer un cambio de paradigma en la identificación de nuevas opciones terapéuticas. Esta estrategia, busca la explotación de aquellos mecanismos específicos de células cancerígenas, que les ayudan a compensar mutaciones críticas para su supervivencia y hacia las cuales desarrollan una dependencia.

La disponibilidad de resultados provenientes de experimentos de silenciamiento de la expresión génica (ya sea mediante shRNA o CRISPR), supone una herramienta muy poderosa para la identificación estas interacciones. Pero aunque estos conjuntos de datos suelen ser públicos, su integración con información genómicos no es trivial.

En este sentido, mediante este análisis, hemos querido proveer de una metodología la detección de vulnerabilidades basándonos en el concepto de letalidad sintética. Se pretende facilitar el análisis de interacciones de letalidad sintética y tiene como finalidad el curado de relaciones con mayor interés biológico.

En cualquier caso, para conseguir nuestro objetivo de integrar esta información a la plataforma PanDrugs, todavía quedan bastantes pasos que habrá que incorporar al análisis.

### **Integración con datos de sensibilidad a drogas en líneas celulares de cáncer.**

Pese a que el par de letalidad sintética haya sido identificado correctamente, no siempre vamos a encontrar drogas capaces de imitar la respuesta sujeta a la inhibición mediante shRNA, ya sea debido a una baja eficacia de los fármacos existentes o por la falta de desarrollo de un fármaco específico.

En este sentido, debemos señalar que sólo un 34% de las interacciones incluidas en SynLethDB disponían de al menos un fármaco capaz de atacar dicha relación. Este bajo porcentaje de genes B para los que existe un tratamiento dirigido, nos da una idea del sesgo existente en la industria farmacéutica hacia el desarrollo de fármacos dirigidos relacionados con el concepto de adicción oncogénica.

Es por ello que es necesario añadir a estos resultados una exploración de los efectos de inhibición de los fármacos atribuidos. El éxito de estos tratamientos farmacéuticos podría suponer un cambio en el marco teórico del diseño de drogas antitumorales.

También es importante notar que gran parte de los fármacos disponibles fueron desarrollados para una enfermedad distinta a cáncer. La letalidad sintética no sólo nos abre el abanico de genes accionables para los cuales podemos desarrollar terapias efectivas, sino que también, nos abre la posibilidad de aplicar reposicionamiento de fármacos para el tratamiento de la enfermedad.

### **Incorporación de datos de silenciamiento de genes mediante CRISPR.**

El portal del proyecto Achilles, ha incorporado recientemente información obtenida mediante el silenciamiento de genes usando CRISPR-Cas9. Como parte de los siguientes pasos del estudio, analizaremos los efectos del silenciamiento total de genes.

Con esta información, esperamos localizar más interacciones de interés. Además, sería interesante explorar cómo varían las consecuencias funcionales según el uso de shRNA o CRISPR, así como su equiparación al silenciamiento de genes mediante inhibición farmacológica.

### **Contexto tumoral.**

Nuestra estrategia de análisis no incorpora la separación por tipo tumoral a la hora de analizar las distintas interacciones. Mediante este acercamiento, buscamos la detección de interacciones de letalidad sintética independientes de su contexto tumoral. La finalidad es localizar aquellas interacciones más fuertes y/o aplicables a un contexto general.

Como paso adicional para nuestro análisis, nos gustaría incluir más líneas celulares provenientes de otros estudios, tanto de silenciamiento como de sensibilidad a droga, con el fin de tener suficiente potencia estadística para estudiar los efectos de estas interacciones dependientes de contexto tumoral.

### **Contexto genético.**

Para intentar mejorar la selección de candidatos, queremos refinar los criterios de clasificación de nuestro filtrado, con el fin de mejorar la discretización de grupos (posiblemente susceptibles y no susceptibles a estrategias de letalidad sintética).

Sería interesante añadir la frecuencia en la que se dan estas mutaciones para cada una de las líneas celulares, así como la información de expresión y metilación disponible, que nos ayudará a añadir un nivel de control extra, localizando aquellos genes reprimidos o expresados.



De hecho, se ha visto que se pueden predecir la mayoría de dependencias utilizando información los valores de expresión de RNA (82%, Tsherniak A *et al.*, 2017), de forma que con este paso esperamos enriquecer nuestros resultados.

### **Redes de interacción.**

Por otro lado, la interpretación de los mecanismos de interacción que llevan a la existencia de pares de letalidad sintética sigue siendo un reto. Es por ello que necesitamos integrar estos datos con redes de interacción proteína-proteína, con el fin de entender los procesos y rutas biológicas que se ven alteradas durante estos eventos de dependencia.

### **SL como biomarcadores de predicción a respuesta a droga - Efecto multitarget.**

Se sabe que la respuesta de las células tumorales a los diversos fármacos a los que es son expuestas, es muy dependiente del contexto genético.

Es por ello que queremos ampliar el contexto de nuestro análisis, pasando de considerar la expresión de exclusivamente los dos genes que conforman el par de interacción de letalidad sintética, a la inclusión de todo el perfil genómico.

Integrando el perfil genómico con el análisis de interacción proteína-proteína, podríamos crear una calificación para indicar la probabilidad de que la línea o tipo tumoral de interés, sea susceptible al tratamiento propuesto. Cuantos más pares de letalidad sintética intercomunicados tengamos, más susceptible será a la utilización de una droga determinada u otra.

### **Integración con PanDrugs.**

Una vez integrada toda la información, queremos crear una categoría de “letalidad sintética” en la base de datos de PanDrugs, en la cual podremos añadir todos los pares de letalidad. Con esta información, buscamos enriquecer la base de datos con información curada de posibles nuevas estrategias, aumentando las opciones terapéuticas disponibles.



## 7. Conclusiones

- La identificación de terapias específicas para las distintas mutaciones o patrones mutacionales, sigue siendo uno de los principales retos para el desarrollo del campo de la medicina de precisión. El concepto de letalidad sintética surge con la finalidad de superar estas dificultades, buscando la explotación de dianas secundarias con el propósito de afectar la viabilidad de la célula tumoral.
- A partir de los datos de mutaciones puntuales y alteraciones en el número de copia, es posible realizar un estudio del estado mutacional de cada una de las líneas celulares provenientes del CCLE. Será necesario incluir los datos de expresión y metilación disponibles, así como separar las líneas celulares según su contexto tumoral.
- Esta información se puede integrar con los datos de esencialidad génica procedentes del proyecto Achilles, con el fin de analizar dependencias secundarias accionables.
- Para el estudio completo de estas dependencias, será necesario añadir una exploración de los efectos de inhibición farmacológica. La letalidad sintética no sólo nos abre el abanico de genes accionables para los cuales podemos desarrollar terapias efectivas, sino que también, nos abre la posibilidad de aplicar reposicionamiento de fármacos para el tratamiento de la enfermedad.
- Añadiendo la estrategia de letalidad sintética a la ya existentes, podemos avanzar en la definición de los subtipos de cáncer y el tratamiento utilizando terapias anticancerígenas. Pacientes que tengan dichas alteraciones pueden ser identificados con exámenes de diagnóstico apropiados, resultando en una subpoblación que tiene una potencial sensibilidad a las correspondientes drogas.



## 8. Bibliografía

Barth BM, Shanmugavelandy SS, Tacelosky DM, Kester M, Morad SA & Cabot MC. Gaucher's disease and cancer: a sphingolipid perspective. *Critical Reviews™ in Oncogenesis*, 2013 18(3).

Bhinder B, Djaballah H; Systematic analysis of RNAi reports identifies dismal commonality at gene-level and reveals an unprecedented enrichment in pooled shRNA screens. *Combinatorial chemistry & high throughput screening*; 2013 16(9), 665-681.

Bridgett S, Campbell J, Lord CJ, Ryan CJ; CancerGD: A Resource for Identifying and Interpreting Genetic Dependencies in Cancer; *Cell Syst.* 2017 Jul 26;5(1):82-86.e3. doi: 10.1016/j.cels.2017.06.002. Epub 2017 Jul 12.

Brunen D, Bernards R; Drug therapy: Exploiting synthetic lethality to improve cancer therapy; *Nat Rev Clin Oncol.* 2017 Jun; 14(6):331-332. doi: 10.1038/nrclinonc.2017.46. Epub 2017 Mar 29.

Campeau PM, Foulkes WD, Tischkowitz MD; Hereditary breast cancer: new genetic developments, new therapeutic avenues; *Hum Genet.* 2008 Aug;124(1):31-42. doi: 10.1007/s00439-008-0529-1. Epub 2008 Jun 25.

Casson L, Howell L, Mathews LA, Ferrer M, Southall N, Guha R, ... & Beverly LJ.. Inhibition of ceramide metabolism sensitizes human leukemia cells to inhibition of BCL2-like proteins. *PLoS One*, 2013; 8(1), e54525.

Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC; Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer; *Proc Natl Acad Sci U S A.* 2011 Jul 26; 108(30):12372-7. doi: 10.1073/pnas.1109363108. Epub 2011 Jul 11.

Cox DA, Fesik SW, Kimmelman AC, Luo J, Der CJ; Drugging the undruggable Ras: mission possible? *Nat Rev Drug Discov.* 2014 Nov; 13(11): 828–851. Published online 2014 Oct 17. doi: 10.1038/nrd4389

Dinu D, Dobre M, Panaitescu E, et al. Prognostic significance of KRAS gene mutations in colorectal cancer - preliminary study. *Journal of Medicine and Life*. 2014;7(4):581-587.

Downward, J; Use of RNA interference libraries to investigate oncogenic signalling in mammalian cells. *Oncogene* 23. November 2004, 8376–8383.

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide; IARC CancerBase No. 11 Lyon, France; *International Agency for Research on Cancer*; 2013.

Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster T, Rahman N, Stratton MR; A census of human cancer genes; *Nature Reviews Cancer* March 2004; 4, 177–183.

García-Parra J, Dalmases A, Morancho B, Arpí O, Menendez S, Sabbaghi M, Zazo S, Chamizo C, Madoz J, Eroles P, Servitja S, Tusquets I, Yelamos J, Lluch A, Arribas J, Rojo F, Rovira A, Albanell J; Poly (ADP-ribose) polymerase inhibition enhances trastuzumab antitumour activity in HER2 overexpressing breast cancer. *Eur J Cancer*. 2014 Oct;50(15):2725-34. doi: 10.1016/j.ejca.2014.07.004. Epub 2014 Aug 12.

Hannun YA, Obeid LM; Principles of bioactive lipid signalling: lessons from sphingolipids; *Nat Rev Mol Cell Biol*. 2008 Feb;9(2):139-50. doi: 10.1038/nrm2329.

Henry B, Möller C, Dimanche-Boitrel MT, Gulbins E, Becker KA; Targeting the ceramide system in cancer. *Cancer Lett*. 2013 May 28;332(2):286-94. doi: 10.1016/j.canlet.2011.07.010. Epub 2011 Jul 23.

Hirsch AJ; The use of RNAi-based screens to identify host proteins involved in viral replication; *Future Microbiol*. 2010 Feb;5(2):303-11. doi: 10.2217/fmb.09.121.

Huang WC, Tsai CC, Chen CL, Chen T Y, Chen YP, Lin YS, ... & Wang CY; Glucosylceramide synthase inhibitor PDMP sensitizes chronic myeloid leukemia T315I mutant to Bcr-Abl inhibitor and cooperatively induces glycogen synthase kinase-3-regulated apoptosis. *The FASEB Journal*, 2011; 25(10), 3661-3673.

Jing Guo, Hui Liu, Jie Zheng; SynLethDB: synthetic lethality database toward discovery of selective and sensitive anticancer drug targets, *Nucleic Acids Research*, Volume 44, Issue D1, 4 January 2016, Pages D1011–D1017, <https://doi.org/10.1093/nar/gkv1108>

Karnoub AE, Weinberg RA; Ras oncogenes: split personalities; *Nat Rev Mol Cell Biol*. 2008 Jul;9(7):517-31. doi: 10.1038/nrm2438.

Kawahara N, Ogawa K, Nagayasu M, Kimura M, Sasaki Y, and Kobayashi H; Candidate synthetic lethality partners to PARP inhibitors in the treatment of ovarian clear cell cancer; *Biomed Rep*. 2017 Nov; 7(5): 391–399. Published online 2017 Sep 27. doi: 10.3892/br.2017.990

Kittler R, Pelletier L, Buchholz F. Systems biology of mammalian cell division; *Cell Cycle*. 2008 Jul 15;7(14):2123-8. Epub 2008 May 21.

Kobayashi N, Nishi T, Hirata T, Kihara A, Sano T, Igarashi Y, & Yamaguchi A; Sphingosine 1-phosphate is released from the cytosol of rat platelets in a carrier-mediated manner. *Journal of lipid research*, 2006; 47(3), 614-621.

Lin NU; Breast cancer brain metastases: new directions in systemic therapy; *Ecancermedicalscience*. 2013; 7: 307. Published online 2013 Apr 18. doi: 10.3332/ecancer.2013.307

Liu J, Beckman BS, Foroozesh M. A review of ceramide analogs as potential anticancer agents. *Future* 2013, 5(12), 1405-1421.

Liu Y, Sun J, Zhao M; ONGene: A literature-based database for human oncogenes. *J Genet Genomics*. 2017 Feb 20;44(2):119-121. doi: 10.1016/j.jgg.2016.12.004. Epub 2016 Dec 26.

Lucci A, Cho WI, Han TY, Giuliano AE, Morton DL, & Cabot, MC . Glucosylceramide: a marker for multiple-drug resistant cancers. *Anticancer research*, 1998, 18(1B), 475-480.

Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, ... & Goodale A. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nature genetics*, 2017; 49(12), 1779.

Morris JP 4th, Wang SC, Hebrok M; KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma; *Nat Rev Cancer*. 2010 Oct;10(10):683-95. doi: 10.1038/nrc2899. Epub 2010 Sep 3.

Munoz DM, Cassiani PJ, Li L, Billy E, Korn JM, Jones MD, Golji J, Ruddy DA, Yu K, McAllister G, DeWeck A, Abramowski D, Wan J, Shirley MD, Neshat SY, Rakiec D, de Beaumont R, Weber O, Kauffmann A, McDonald ER 3rd, Keen N, Hofmann F, Sellers WR, Schmelzle T, Stegmeier F, Schlabach MR; CRISPR Screens Provide a Comprehensive Assessment of Cancer Vulnerabilities but Generate False-Positive Hits for Highly Amplified Genomic Regions. *Cancer Discov.* 2016 Aug; 6(8):900-13. doi: 10.1158/2159-8290.CD-16-0178. Epub 2016 Jun 3.

Nicholson KM, Quinn DM, Kellett G L, & Warr JR; Preferential killing of multidrug-resistant KB cells by inhibitors of glucosylceramide synthase. *British journal of cancer*, (1999); 81(3), 423-430.

Nowsheen S, Cooper T, Stanley JA, and Yang ES; Synthetic Lethal Interactions between EGFR and PARP Inhibition in Human Triple Negative Breast Cancer Cells; *PLoS One*. 2012; 7(10): e46614. Published online 2012 Oct 11. doi: 10.1371/journal.pone.0046614

O'Neil NJ, Bailey ML, Hieter P. Synthetic lethality and cancer; *Nat Rev Genet.* 2017 Oct;18(10):613-623. doi: 10.1038/nrg.2017.47. Epub 2017 Jun 26.

Pal T, Permuth-Wey J, Betts JA, Krischer JP, Fiorica J, Arango H, LaPolla J, Hoffman M, Martino MA, Wakeley K, Wilbanks G, Nicosia S, Cantor A, Sutphen R; BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases; *Cancer*. 2005 Dec 15;104(12):2807-16.

Pitson SM; Regulation of sphingosine kinase and sphingolipid signaling; *Trends Biochem Sci.* 2011 Feb;36(2):97-107. doi: 10.1016/j.tibs.2010.08.001. Epub 2010 Oct 1.

Prahallad A, Sun C, Huang S, Di Nicolantonio F, Salazar R, Zecchin D, Beijersbergen RL, Bardelli A, Bernards R; Unresponsiveness of colon cancer to BRAF (V600E) inhibition through feedback activation of EGFR. *Nature*. 2012; 483, 100–103.

Roberts PJ, Der CJ, Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer; *Oncogene*. 2007 May 14;26(22):3291-310.

Rosen H, Goetzl EJ; Sphingosine 1-phosphate and its receptors: an autocrine and paracrine network. *Nature Reviews Immunology*, 2005 5(7), 560-570.



Ryan CJ, Lord CJ, Ashworth A; DAISY: Picking Synthetic Lethals from Cancer Genomes *Cancer Cell* 2014 Sep 8;26(3):306-308. doi: 10.1016/j.ccr.2014.08.008.

Scholl C, Fröhling S, Dunn IF, Schinzel AC, Barbie DA, Kim SY, Silver SJ, Tamayo P, Wadlow RC, Ramaswamy S, Döhner K, Bullinger L, Sandy P, Boehm JS, Root DE, Jacks T, Hahn WC, Gilliland DG; Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells; *Cell* 2009 May 29;137(5):821–834. doi: 10.1016/j.cell.2009.03.017.

Shah NP, Tran C, Lee FY, Chen P, Norris D, Sawyers CL; Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science*; 2004 Jul 16;305(5682):399-401.

Shida D, Fang X, Kordula T, Takabe K, Lepine S, Alvarez SE, Milstien S, Spiegel S; Cross-talk between LPA1 and epidermal growth factor receptors mediates up-regulation of sphingosine kinase 1 to promote gastric cancer cell motility and invasion. *Cancer Res.* 2008. 68, 6569–6577.

Sinha S, Thomas D, Chan S, Gao Y, Brunen D, Torabi D, Reinisch A, Hernandez D, Chan A, Rankin EB, Bernards R, Majeti R, Dill DL; Systematic discovery of mutation-specific synthetic lethals by mining pan-cancer human primary tumor data; *Nature Communications* 8, Article number: 15580 (2017); doi:10.1038/ncomms15580

Tsherniak A, Vazquez F, Hahn WC; Defining a Cancer Dependency Map, *Cell*, 2017 Jul 27;170(3):564-576.e16. doi: 10.1016/j.cell.2017.06.010

Weidle UH, Maisel D, Eick D; Synthetic lethality-based targets for discovery of new cancer therapeutics; *Cancer Genomics Proteomics*. 2011 Jul-Aug;8(4):159-71.

Westbrook TF, Martin ES, Schlabach MR, Leng Y, Liang AC, Feng B, Zhao JJ, Roberts TM, Mandel G, Hannon GJ, et al. 2005. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121:837–848

Xu Y, Shen Z, Wiper DW, Wu M, Morton RE, Elson P, Kennedy AW, Belinson J, Markman M, Casey G. Lysophosphatidic acid as a potential biomarker for ovarian and other gynecologic cancers. *JAMA*; 1998 280, 719–723.

Yersal O and Barutca S; Biological subtypes of breast cancer: Prognostic and therapeutic implications; *World J Clin Oncol*. 2014 Aug 10; 5(3): 412–424. Published online 2014 Aug 10. doi: 10.5306/wjco.v5.i3.412

## 9. Anexos

### Variantes genómicas

Una variante genómica es un tipo de variación alélica del genoma que aparece como consecuencia de una mutación y que es de carácter persistente, apareciendo en miembros de una misma población o poblaciones.

Existen distintas clasificaciones para la caracterización de las variantes genómicas.

En primer lugar, basándonos en su **origen**, podemos hablar de variantes de tipo germinal y de variantes somáticas. Las primeras, se encuentran en todas las células del organismo, son de carácter hereditario y pueden asociarse a la predisposición a desarrollar ciertos tipos de enfermedades. Mientras que las segundas, sólo están presentes en células somáticas y se adquieren a lo largo de la vida del individuo.

Según su **tamaño**, las variantes se pueden clasificar en:

1. SNPs (*single nucleotide polymorphisms*), alteraciones que afectan a un sólo nucleótido en la secuencia del ADN. Estas variaciones deben darse al menos en un 1% de la población para ser considerada como un SNP.
2. Indels, polimorfismos que abarcan segmentos de ADN menores a 1kb.
3. Variantes estructurales, aquellas cuyo tamaño es superior a 1kb. Dentro esta clasificación, encontramos los CNVs o *copy number variants*. Este tipo de variantes pueden llegar a alterar la dosis génica y los patrones de expresión de la célula y están relacionadas con diversas enfermedades o un aumento en la predisposición a su desarrollo.

Y en el caso de que aparezcan en regiones codificantes, según su **efecto**:

1. Mutación con desplazamiento de marco de lectura: aquella que es capaz de cambiar el patrón de traducción y por tanto, el producto proteico de ese gen.
2. Mutación de cambio de sentido: se provoca un cambio en la secuencia de aminoácidos.
3. Mutación sin sentido: supone la aparición de un codón que provoca la parada de la traducción, de forma que la proteína final queda truncada.
4. Mutaciones silenciosas: no supone ningún cambio en el producto proteico.

### **Definición de diana directa o gen biomarcador en cuanto a la respuesta a drogas**

En cada asociación del tipo gen-droga, cada gen puede tener un rol diferente en la respuesta al fármaco. Definimos como **diana directa**, a aquel gen que contribuye al fenotipo de la enfermedad y está directamente afectado por el fármaco (molécula pequeña, anticuerpo monoclonal...)

Denominamos **biomarcador** a un gen cuyo estatus genético está asociado a la respuesta a droga, pero cuyo producto proteico no es la diana directa de la droga.

### **Genomics of Drug Sensitivity of Cancer (GDSC)**

El proyecto GDSC (<http://www.cancerrxgene.org/>) consiste en un conjunto de datos público para el estudio de la respuesta a diferentes fármacos actualmente en uso, en líneas celulares tumorales. Para el estudio, se han caracterizado molecularmente más de 1000 líneas celulares con distintos orígenes tumorales. Entre los compuestos analizados, se incluyen tanto quimioterapias como terapias dirigidas.

### **Reposicionamiento de fármacos**

El reposicionamiento de fármacos consiste en la investigación y atribución de nuevas indicaciones en fármacos previamente aprobados.